<table>
<thead>
<tr>
<th>Title</th>
<th>Structural basis for promiscuity and specificity during Candida glabrata invasion of host epithelia.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>2012</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10220/12333">http://hdl.handle.net/10220/12333</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 2012 National Academy of Sciences. This paper was published in Proceedings of the National Academy of Sciences and is made available as an electronic reprint (preprint) with permission of National Academy of Sciences. The paper can be found at the following official DOI: [<a href="http://dx.doi.org/10.1073/pnas.1207653109">http://dx.doi.org/10.1073/pnas.1207653109</a>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.</td>
</tr>
</tbody>
</table>
Structural basis for promiscuity and specificity during *Candida glabrata* invasion of host epithelia

Manuel Maestre-Reyna, Rike Diderrich, Maik Stefan Veelders, Georg Eulenburg, Vitali Kalugin, Stefan Brückner, Petra Keller, Steffen Rupp, Hans-Ulrich Mösch, and Lars-Oliver Essen

The human pathogenic yeast *Candida glabrata* harbors more than 20 surface-exposed, epithelial adhesins (Epas) for host cell adhesion. The Epa family recognizes host glycans and discriminates between target tissues by their adhesin (A) domains, but a detailed structural basis for ligand-binding specificity of Epa proteins has been lacking so far. In this study, we provide high-resolution crystal structures of the Epa1A domain in complex with different carbohydrate ligands that reveal how host cell mucin-type O-glycans are recognized and allow a structure-guided classification of the Epa family into specific subtypes. Further detailed structural and functional characterization of subtype-switched Epa1 variants shows that specificity is governed by two inner loops, CBL1 and CBL2, involved in calcium binding as well as by three outer loops, L1, L2, and L3. In summary, our study provides the structural basis for promiscuity and specificity of Epa adhesins, which might further contribute to developing anti-adhesive antimycotics and combating *Candida* colonization.

molecular recognition | T-antigen | lectin | fungal pathogen

The strict mammalian commensal yeast *Candida glabrata* can act as a health care associated opportunistic pathogen. In immune-compromised patients, colonization results in lethal systemic infections and sepsis (2, 3). Besides that, *C. glabrata* is the causative agent of several less severe, mucosa-related, oral or urogenital diseases (4, 5). Due to its resistance against traditional antimycotics (6), this pathogen is considered to be highly emergent (4), and new therapeutic tools are being actively sought. One putative target for drug development is the epithelial adhesin (Epa) machinery, which is represented in *C. glabrata* by up to 23 different genes. Apart from *EPA1*, which confers adhesion of fungal cells to mammalian surfaces in vitro (7) and in vivo (8), the function of other EPA genes is only partly understood. For example, *EPA6* and *EPA7* have been identified as virulence factors, and *EPA6* is further linked to biofilm formation and colonization of the vaginal epithelium (6). Interestingly, most of the EPA4 genes are clustered in subtelomeric regions, indicating rapid genetic adaptation to varying environmental conditions during host colonization (1). Also, expression of *EPA1*, *EPA6*, and *EPA7* is regulated by a distinct silencing mechanism via the Sir machinery and mutation of *SIR* genes causes hypervirulence and enhanced kidney colonization (9).

Members of the Epa family are highly glycosylated mannoproteins with a modular architecture common to fungal GPI-anchored cell wall proteins (GPI-CWP) (10). While the N-terminal region (adhesion, or A domain) resembles C-type lectins mediating calcium-dependent adhesion to carbohydrate structures, the central serine- and threonine-rich segment (B domain) comprises a variable number of highly O-glycosylated repeats. Finally, a C-terminal anchor (C domain) cross-links the adhesin covalently to the cell wall (7, 11). Epa A domains are distantly related to flocculins from *Saccharomyces cerevisiae*, which also harbor A domains belonging to the PA14/Flo5-like family (12, 13). For the A domain of the flocculin Flo5, mannoside recognition involves a novel motif not found in other C-type lectins (12), the DcsID motif, whose role in galactose recognition by Epa1 has been recently confirmed (14). Other structural features determining sugar specificities by Flo5-like domains are so far unclear and do not allow correlating the different Epa proteins to their distinct and mostly still unknown specificities.

Recent, semiquantitative functional characterizations of Epa1, Epa6, and Epa7 showed that the specificity of these three adhesins is related (15). All three adhesins require a terminal galactose containing carbohydrate ligand for host recognition. Epa6 binds to oligosaccharides comprising galactose linked via 1–3 or 1–4 glycosidic bonds to glucose, galactose, or their N-acetylated derivatives. Furthermore, Epa6 does not discriminate between α- and β-glycosidic linkages. In contrast, Epa1 has a narrower specificity, mainly excluding α-linked carbohydrates. Epa7 is the most specific of the three adhesins and almost exclusively binds to lactose (Galβ1–4Glc), Galβ1–3Gal, or their respective N-acetylated derivatives (15). Finally, adhesion specificity of EpaA domains is claimed to be governed by hypervariable stretches comprising amino acids 114–139 and 227–231 (15). However, a structural basis correlating these findings to Epa specificities has not been established.

Here, we present high-resolution crystal structures of the Epa1A domain complexed to cognate disaccharide ligands including Galβ1–3Glc or the T-antigen (Galβ1–3GalNAc). We further provide a structure-guided classification of the Epa family into specific subtypes and present structures of subtype-switched Epa1 variants with modified binding sites that correspond to Epa2, Epa3, and Epa6. Together with high-throughput carbohydrate-binding screens and fluorescence titration analyses these structures uncover the critical differences within the Epa protein family in regard to specificity and promiscuity. In vivo, subtype-specific binding of Epa proteins to epithelial surfaces in part depends on the specificity-determining loop region CBL2. Overall, *Candida* species can generate a large repertoire of Epa adhesins for colonizing various host niches by exerting subtle structural changes to a common motif.

Results and Discussion

**Epa1A Preferentially Recognizes Galβ1–3 Glycans.** To determine the carbohydrate specificity profile of Epa1, the recombinant, fluorescence-labeled Epa1A domain was analyzed by glycan array screening [ Consortium for Functional Glycomics (CFG)]. Binding of Epa1A to such arrays at 200 µg/mL showed a marked


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, wwwpdb.org (PDB ID codes 4AF9, 4AF4, 4AF8, and 4AC6).

*To whom correspondence may be addressed. E-mail: essen@chemie.uni-marburg.de or moesch@staff.uni-marburg.de.*

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207653109/-/DCSupplemental.
preference for glycans containing galactose at nonreducing ends (Fig. 1A and B). Both α- and β-linked galactoses are bound by Epa1A, but the binding capability for Galβ1-terminated glycans is significantly higher than for those with Galα-terminal residues. Previous binding studies with Epa1A presented at the surface of S. cerevisiae indicated that this adhesin does not discriminate between mucin-relevant Galβ1-3 and Galβ1-4 disaccharides (15). Here, also no obvious preference was found when comparing Galβ1-3 and Galβ1-4-containing glycans, but all of the best binders within the Galβ1-4 group contain terminal Galβ1-3 branches (Fig. 1B). To further discriminate between stringent primary and nonstringent secondary specificities, binding at lower protein concentrations was conducted. At 50 μg/mL Epa1A, a moderate decrease in the group of Galβ1-4-containing glycans was observed that sharply drops at 1 μg/mL. Likewise, oligosaccharides with terminal α-linked galactose moieties show diminished binding at restricting Epa1A concentrations, further pointing to a preference for β-linkage. Best binders are Galβ1-3 group members, especially those with a Galβ1-3GalNAc motif including the T-antigen (glycans 137 and 140, Fig. 1A). In contrast, β-4-linked lactosides, which are known to inhibit Candida adherence to host cells in the millimolar range (7), show only weak (lactose, glycans 169 and 170) or very moderate (N-acetyllactosamine, glycans 167 and 168) binding to Epa1A. In summary, our data provide a refined specificity profile for Epa1A and offer compelling evidence that this adhesin prefers Galβ1-3 disaccharides, but lacks high affinity for Galβ1-4 disaccharides. This conclusion is further supported by our fluorescence titration analysis (Fig. S1), which demonstrates that Epa1A has at least a 16-fold higher binding affinity for the T-antigen (K_D = 2.1 ± 0.3 μM, Galβ1-3GalNAc) than for milk-derived lactose (34.6 ± 6.1 μM), which is likely to contain significant amounts of a Galβ1-3Glc contaminant as delineated by our structural analysis below.

**Overall Epa1A Structure.** To determine the structure of Epa1A, crystals were obtained at low pH, but only in the presence of high concentrations of lactose. The Epa1A crystal structure was solved by molecular replacement, using a truncated homology model based on the S. cerevisiae Flo5A structure (12). Despite a pairwise sequence identity of less than 25%, the resulting structure (Tables S1 and S2) reveals a high similarity with Flo5A, as indicated by an rmsd of 1.4 Å over 121 common C_C atoms (Fig. 24). The barrel-shaped Epa1A domain has an overall dimension of 57 × 42 × 36 Å³, consists of 15 β-strands, and is composed via a calcium ion to its respective disaccharide ligand (Fig. 24). Eleven of the strands form an antiparallel β-sandwich motif with PA14/Flo5-like topology (16). An L-shaped stretch, composed of the remaining four β-strands and the N and C termini, wraps around the β-sandwich and protects its bottom end from solvent access (12). This end of Epa1A is adjacent to the repetitive stalk of the Epa1 B domain. The two disulfide bridges CS50-C179 and C180-C263, respectively, which covalently tether the N and C termini to the β-sandwich domain (Fig. 24), are preserved in Flo5, suggesting that this rigid linkage between A and B domains is a general feature of GPI-CWP adhesins.

Which structural motifs define the ligand-binding site in Epa1A? In contrast to S. cerevisiae flocculins, Epa1A lacks the Flo5-like subdomain, which in budding yeast confers specific interactions with terminal α-1,2-mannobioside ligands (12). Surprisingly, the difference electron density map within the binding pocket of the Epa1A/lactose cocrystals clearly ruled out lactose, i.e., Galβ1-4Glc, as bound ligand. At 1.5 Å resolution only a Galβ1-3Glc disaccharide, a contaminant from commercially available lactose, was consistent with the electron density observed (Fig. 2B). This serendipitous finding corroborates the strong preference of Epa1A for Galβ1-3 over excess of Galβ1-4-linked disaccharides as indicated by our glycan profiling and

---

**Fig. 1.** Glycan binding by the Epa1A domain. (A) Binding profiles of Epa1A at different protein concentrations (1, 50, and 200 μg/mL), using the CFG array V4.1 harboring 451 different glycan structures. Relative fluorescence units as monitored reflect relative affinities toward the corresponding glycan. Glycans bound by Epa1A are indicated by their CFG array numbers. Overall, Epa1A recognizes exclusively terminal galactosides. Green, Galα group; red, Galβ1-3 group; blue, Galβ1-4 group. (B) Groups that are recognized by Epa1A are shown; their locations within the CFG array are given in parentheses. Structural formulas are described according to the CFG nomenclature. At 200 μg/mL, Epa1A recognizes both Galβ1-3 (red) and Galβ1-4 (blue) glycans. However, the latter can harbor terminal β1-3 ramifications (red dotted lines). At low concentrations, Epa1A has a much narrower specificity profile.
fluorescence titration analyses. The binding site of the Galβ1–3
disaccharide is formed of five loops: two inner, short ones in-
volved in calcium binding and ligand recognition (CBL1 and
CBL2) and three outer, longer loops (L1, L2, and L3), which
shield CBL1 and CBL2 from solvent access (Fig. 2D). CBL1
links β-strand 8 with 9 and is hallmarked by an unusual cis-
peptide between D165 and D166, the DcisD motif. This motif
is crucial for the C-type lectin function of Flo5A (12) and is con-
served throughout the entire Epa family (Fig. S2). CBL2 links
β-strands 12 and 13 and directly contributes to specific glycan
binding by four of its side chains, R226, E227, Y228, and D229 (Fig. 2 D and E), which we further refer to as CBL2 positions I– IV (Fig. 3). Finally, the Caβ4-binding site comprises the side
chain of N225, the carbonyl groups of peptide bonds within
CBL2, and the carboxylic side chains of D165 and D166, the
DcisD motif of CBL1. Moreover, the Caβ4 ion directly interacts
with the nonreducing galactose moiety over the 3- and 4-hydroxyls
formed by C78 and C119 cross-links L1 and L2 to arrest their
mobility, whereas the side chains from the CBL2 tip interact with glucose or GalNAc hydroxyls.

Disaccharide Recognition by the Inner and Outer Subsites of Epa1A.
What is the structural basis for disaccharide discrimination by Epa1A? The Epa1A-Galβ1–3Glc complex indicates that at least
two subsites are crucial for glycan recognition: (i) the inner
subsite, which is formed of CBL1, the Caβ4 ion, R226 from
CBL2, and W198 from loop L3 and confers specific binding to the
terminal galactose, and (ii) the outer subsite, which consists of
the side chains of CBL2 together with loops L1, L2, and L3 and
interacts with the peripheral hexose moiety (Fig. 2A and C).
Moreover, loop L1 is not part of the core PA14/Flo5A-like do-
main, but connects β-strands 2 and 3. Also, a disulfide bridge
formed by C78 and C119 cross-links L1 and L2 to arrest their
conformation for shielding the binding site from bulk solvent.

The Epa1A-Galβ1–3Glc complex further reveals that the gal-
actose moiety in the inner subsite adopts a flipped orientation,
which differs from the orientation of mannose in Flo5A (12). This
flipped binding mode is crucial for cofiguring the pyranose moiety
with residue W198 from loop L3 as well as for hydrogen bonding
between the R226 side chain and the 2- and 3-hydroxyls of gal-
actose (Fig. 2D). This suggests that the indole side chain of W198
together with R226 is crucial for selecting the terminal galactose
(Figs. S3 and S4), a conclusion that is supported by our finding that
an Epa1 W198A variant shows only marginal binding activity (Fig.
S5). We also performed molecular dynamics studies for ligand
binding by Epa1A (Movies S1, S2, and S3), which indicate that
W198 may exert an organizing effect on loop L3, because they
show that the stretch W198–T202 of loop L3 adopts the observed
degree of order only upon complexation of galactose.

The outer subsite apparently selects the type of glycosidic
bondage and nature of the hexose moiety linked to the terminal
galactose. The Epa1A-Galβ1–3Glc structure shows that the site
for recognition of Galβ1–3-linked hexoses not only comprises
E227 and Y228 of CBL2 (Fig. 2 B and D), but also loops L1 and
L2 lining the outer site. The glutamic acid side chain projected
from position II forms hydrogen bonds with the 2-hydroxyls of
both hexoses of Galβ1–3Glc as well as with loop L2 via the
peptide group of G118–C119. Likewise, the L2 loop wraps with
A115–G118 around Y228 and thereby stabilizes its packing with
the pyranose ring of the glucose within the Galβ1–3Glc ligand.

To obtain further insights into disaccharide binding specificity,
the structure of the cognate Epa1A-T-antigen complex was solved
at 1.24 Å resolution. This structure reveals the same mode of glycan
recognition at its inner subsite as found for the Epa1A-Galβ1–3Glc
complex, but major differences at the outer subsite (Fig. 2 C and
E). Whereas in the Epa1A-Galβ1–3Glc complex the glucose moiety
is coplanar to galactose, the N-acetylgalactosamine moiety of the
T-antigen adopts an orthogonal conformation when complexed to
Epa1A. This diverging conformation is apparently caused by the
hydrogen bond between the axial 4-hydroxyl group and E227 of
the Epa1A binding site is highly specific for disfavoring Galβ1–4-linked glycans. Molecular modeling of a lactose
complex and subsequent molecular dynamics studies show that
a β-1 linkage fails to position the second hexose in a defined
conformation within the outer subsite, e.g., by packing with Y228
(Fig. S4 and Movies S3 and S4). Nevertheless, the outer subsite
harbors sufficient flexibility to allow different binding modes, as
shown for the T-antigen. This suggests that EpaA domains permit
a certain degree of promiscuity for disaccharide binding in their
outer, but not in their inner ligand-binding subsite.

Differential Promiscuity of Epa Subtypes Is Dominated by CBL2.
We next generated a sequence alignment of 19 EpaA domains that is
based on the structures of Epa1A and Flo5A (Fig. S2) and provides a structure-guided phylogenetic tree of the Epa family
(Fig. 3). Four prominent branches appear in this tree, which cluster 11 EpaA domains into four subtypes, which were named
Epa1 (Epa1A, Epa6A, Epa7A), Epa2 (Epa2A, Epa4A, Epa5A, Epa19A), Epa3 (Epa3A, Epa22A), and Epa9 (Epa9A, Epa10A), respectively. Remarkably, the CBL2 positions I–IV are mostly
sufficient to discriminate between the different subtypes (Fig.
3B). For Epa subtype members, this motif consists of REYD or
RDND, and the Epa2 and Epa9 subtype adhesins contain rel-
ated RDNN or RDYH motifs. Only Epa3 subtype proteins show a highly diverging IGKD motif. In addition, the Epa9 subtype
differs from others by a prominent elongation within the L1 loop.
To explore how CBL2 affects the specificity of different Epa
subtype proteins, the CBL2 motif of Epa1A was exchanged for the
motif of Epa2A and Epa3A (Epa1→2A and Epa1→3A variants). Because the Epa1 subtype includes two different CBL2
signatures, an Epa1→6A variant was also generated. The Epa9 subtype was not further investigated due to its di-
verging L1 loop. In the next step, fluorescence titration analysis
of the subtype-switched Epa1A variants was performed. This
analysis revealed that the binding affinities of the Epa1→6A
variant for lactose and the T-antigen resemble those of Epa1A

Fig. 2. Glycan-binding site of the Epa1 adhesin. (A) Structure of the Epa1A domain (gray) with Galβ1–3Glc (Gal, yellow; Glc, blue). The outer subsite for glycan binding is formed by the CBL loops (cyan) and a Caβ4 ion (orange). The Flo5A structure is shown as a green transparent overlay, for comparison purposes. The Galβ1–3Glc (β and α) ligands are shown within the Epa1A-binding site with their SIGMAA-weighted omit electron density maps, respectively (contouring level 1σ for Cβ, 0.3 electrons/Å̃), and (D) Perpendicular views on the Epa1A-binding site either with the Galβ1–3Glc ligand (D) or with the T-antigen (E). Loops L1 and L2 are covalently linked via a cysteine bridge. W198 from L3 and Caβ4 interact directly with the galactose moiety, whereas the side chains from the CBL2 tip interact with glucose or GalNAc hydroxyls.

16866 | www.pnas.org/cgi/doi/10.1073/pnas.1207653109
Maestre-Reyna et al.
Fig. 3. Phylogeny of C. glabrata EpaA domains. (A) Phylogenetic tree of EpaA domains. The tree is based on a structure-guided multiple-sequence alignment (Fig. S2), using the crystal structures of C. glabrata Epa1A and S. cerevisiae Flo5A (PDB ID: 2XJP). The Epa family shows four major, distinct subtypes: Epa1 (gray), Epa2 (green), Epa3 (blue), and Epa9 (orange). The scale bar indicates phylogenetic distances in number of amino acid substitutions per position. (B) Alignment of the CBL2 regions (cyan box). The Epa1, Epa2, Epa3, and Epa9 subtypes are color-shaded as in A. The Roman numerals indicate the four variable residues within CBL2.

\( K_D = 33.5 \pm 6.6 \text{ mM} \) and \( 1.7 \pm 0.4 \text{ mM} \). In contrast, the Epa1→2A variant has a fourfold lower affinity for lactose (\( K_D = 128 \pm 16.3 \text{ mM} \)). Unfortunately, this variant rapidly quantifies of binding despite its obvious interaction with this glycan. For the Epa1→3A variant, we detected only extremely low affinity by this variant, similar to Epa1 (Fig. 4). For Epa1→2A and Epa1→6A variants were found to harbor the Gal_β1→3Glc complex (Fig. 4). This suggests that the changes at positions II and III in CBL2 (E227D, Y228N) cause a wider and more flexible outer subsite when compared to that in Epa1A (Fig. 4). As a consequence, the reducing end of the di-galactoside (Fig. 4). In contrast to Epa1, we found strongly diminished binding compared to Epa1A and Epa1→6A, which were highly biased toward glycans with terminal, α-linked galactose moieties. In summary, all variants analyzed by glycan array profiling are highly specific for galactose in the inner subsite, whereas the outer subsite is promiscuous, being occupied by glucose, N-acetylgalcosamine, galactose, or N-acetylgalactosamine (Figs. S6 and S7), a finding that explains previous observations (15).

Next, the adhesion behavior of different Epa subtypes was analyzed in vivo, using a heterologous S. cerevisiae system to reintegrate EpaA domains on the cell surface (Fig. S8) and the human colorectal, epithelial cell line Caco-2. The T-antigen is a typical mammalian surface glycan, which is present at high densities on Caco-2 cell surfaces, but not on fungal cells. As expected, S. cerevisiae strains that lack an EpaA domain failed to adhere to epithelial cells (Fig. 5A). In contrast, strains presenting Epa1A or subtype-switched Epa1→6A strongly adhere to Caco-2 cell layers. Obviously, T-antigen–presenting epithelial cells are well recognized by the Epa1 subtype independently of the used sequence of its CBL2 motif (Fig. 5A). Furthermore, the 15-fold lower affinity of Epa1→3A to recognize the T-antigen in vitro is mirrored in vivo, because strains presenting either Epa1→3A or native Epa3 only inefficiently bind to Caco-2 cells (Fig. 5A). Epa1→2A that closely mediates binding to Caco-2 cells presents a different scenario. Given an only fourfold reduced in vitro binding affinity to lactose compared to Epa1A and Epa1→6A (Fig. S1), one may expect that the affinity or at least \( k_{off} \) of Epa1→2A against the T-antigen is similar to that of the Epa1A subtypes. However, the native Epa2A domain fails to confer significant epithelial cell binding, suggesting that in this case switching the specificity and binding characteristics between Epa subtypes seems to require more than exchanging the CBL2 motif. Additional structural features of the outer subsite like the L1 and L2 loops may hence contribute to binding and specificity. Nevertheless, our results indicate that the CBL2 loop plays a prominent role in conferring specificity and promiscuity to Epa family adhesins and thereby significantly affects efficient host cell recognition.

Subtle Structural Changes Cause Promiscuity. To obtain further detailed insights into Epa protein functionality, the structures of the subtype-switched Epa1A variants were determined by using the orthorhombic crystal form obtained from cocystalization with lactose. This analysis revealed structural alterations of the Epa1A variants, which were restricted to their glycan-binding sites when compared to Epa1 (Fig. 4). Similar to Epa1, the Epa1→2A and Epa1→6A variants were found to harbor the Gal_β1→3-linked di-saccharide with a well-defined positioning of the galactose moiety in the inner subsite. In contrast to Epa1, however, the terminal glucose residue was present in two alternative conformations within the outer subsite (Fig. 4A). This suggests that the change at positions II and III in CBL2 (E227D, Y228N) cannot be easily accommodated in Epa6-like binding sites, because strains presenting either Epa1→3A or native Epa3 only inefficiently bind to Caco-2 cells (Fig. 5A). Epa1→2A that closely mediates binding to Caco-2 cells presents a different scenario. Given an only fourfold reduced in vitro binding affinity to lactose compared to Epa1A and Epa1→6A (Fig. S1), one may expect that the affinity or at least \( K_D \) of Epa1→2A against the T-antigen is similar to that of the Epa1A subtypes. However, the native Epa2A domain fails to confer significant epithelial cell binding, suggesting that in this case switching the specificity and binding characteristics between Epa subtypes seems to require more than exchanging the CBL2 motif. Additional structural features of the outer subsite like the L1 and L2 loops may hence contribute to binding and specificity. Nevertheless, our results indicate that the CBL2 loop plays a prominent role in conferring specificity and promiscuity to Epa family adhesins and thereby significantly affects efficient host cell recognition.

Subtle Structural Changes Cause Promiscuity. To obtain further detailed insights into Epa protein functionality, the structures of the subtype-switched Epa1A variants were determined by using the orthorhombic crystal form obtained from cocystalization with lactose. This analysis revealed structural alterations of the Epa1A variants, which were restricted to their glycan-binding sites when compared to Epa1 (Fig. 4). Similar to Epa1, the Epa1→2A and Epa1→6A variants were found to harbor the Gal_β1→3-linked di-saccharide with a well-defined positioning of the galactose moiety in the inner subsite. In contrast to Epa1, however, the terminal glucose residue was present in two alternative conformations within the outer subsite (Fig. 4A). This suggests that the change at positions II and III in CBL2 (E227D, Y228N) cannot be easily accommodated in Epa6-like binding sites, because strains presenting either Epa1→3A or native Epa3 only inefficiently bind to Caco-2 cells (Fig. 5A). Epa1→2A that closely mediates binding to Caco-2 cells presents a different scenario. Given an only fourfold reduced in vitro binding affinity to lactose compared to Epa1A and Epa1→6A (Fig. S1), one may expect that the affinity or at least \( K_D \) of Epa1→2A against the T-antigen is similar to that of the Epa1A subtypes. However, the native Epa2A domain fails to confer significant epithelial cell binding, suggesting that in this case switching the specificity and binding characteristics between Epa subtypes seems to require more than exchanging the CBL2 motif. Additional structural features of the outer subsite like the L1 and L2 loops may hence contribute to binding and specificity. Nevertheless, our results indicate that the CBL2 loop plays a prominent role in conferring specificity and promiscuity to Epa family adhesins and thereby significantly affects efficient host cell recognition.

In the Epa1→2A–Galβ1→3Glc complex, the glucose adopts an orthogonal conformation similar to that of the Epa1→6A–Galβ1→3Glc complex (Fig. 4). Having the same exchanges at CBL2 positions II and III, the Epa1→2A variant recognizes Galβ1→3Glc joined glycans as well. In addition, the change in position IV (D229N) reduces the negative charge next to the galactose ligand. As a consequence, Epa1→2A exhibits a marked preference for terminal galactosides that are sulfated at the 6-hydroxyl position (Fig. 5B). These glycan structures are prevalent in sulfomucins, which are in turn predominant components of...
intestinal mucosa (20). The structure of the Ep1→2A-Galβ1–3Glc complex offers sufficient room in the inner subsite to position the 6-sulfate of these terminal sulfo-galactosides close to N229. In contrast, the Ep1 subtype is hindered to recognize sulfo-galactosides due to possible electrostatic repulsion by N229. In contrast, the Epa1 subtype is hindered to recognize sulfo-galactosides close to the 6-sulfate of these terminal sulfo-galactosides. The lower fluorescence signals of Ep1→2A and Ep1→3A indicate weaker binding to glyans than that of Ep1A and Ep1→6A. (Lower) Top view of subtype-switched Ep1A-binding sites with SIGMAA-weighted omit electron densities for ligands complexed to Ep1→6 (anthracite gray, contouring level: 0.8σ, 0.29electrons/Å³), Ep1→3A (blue, 0.8σ, 0.28electrons/Å³) and Ep1→2A (green, 0.6σ, 0.19electrons/Å³). The orientation of the disaccharide within the binding pocket is shown (Upper Right Inset). CBL loops of the inner subsite are marked in cyan and L loops of the outer subsite in orange.

**Conclusions**

Our detailed structural and functional characterization of Ep1A and subtype-switched variants demonstrates that the specificity of C. glabrata Ep1A variants for different glycan structures is exerted by a distinct region of their A domain. This region comprises the calcium-coordinating loops CBL1 and CBL2 as well as parts of loop L3 and bears modular characteristics to generate Ep1 subtype variability. The specificity profiles of Ep1A and its variants strengthen and refine the link between these C. glabrata epithelial adhesins and the cores of mucin-type O-glycans (15). Importantly, we found that the T-antigen, which represents the core 1 of mucin-type O-glycans (21) and is part of most mucin-derived glycans (22), is one of the best ligands for all Ep1A variants investigated in this study. Other glycans that we found to be recognized by Ep1A and variants are also related to mucins and...
include other mucin-type O-glycan cores or N-glycans, e.g., sulfogalactose–containing glycans.

Our study further reveals that the structural basis for the specificity and promiscuity in Epa–glycan interactions is significantly affected by the properties of CBL2. This finding is consistent with previous domain-swapping experiments between Epa6 and Epa17, which identified a larger, CBL2-covering region for controlling specificity (15). Specificity of the inner sub-site for galactose is dictated by recognition of the cis-diol formed by the 3- and 4-hydroxyls. Furthermore, position I of CBL2 is occupied by an arginine in all Epa subtypes except Epa3. This crucial residue forms two hydrogen bonds with the 2-hydroxyl of the bound galactose and packing to loop L1 via W79 and Y83 stabilizes its conformation. In the outer sub-site, position III of CBL2 presents either large aromatic residues (Y, Epa1, Epa9, Epa10, Epa12, Epa15, and Epa20; F, Epa23) or polar amino acids like asparagine (Epa2, Epa4, Epa5, Epa6, and Epa19). Whereas the former promote packing with planar glucose/galactose derivatives, resulting in specific interactions with either terminal Galβ1–3Glc(NAc) or Galβ1–3Gal(NAc), the latter can increase promiscuity (Epa6) or enable additional bindings, e.g., sulfated galactoses (Lys: Epa3, Epa22). Position IV appears to mainly control the degree of allowed modification in the terminal galactose moiety. Together, CBL2 positions I–IV may be proposed to form a simple structural code for Epa specificity. However, the differences found for binding to host cell monolayers by Epa1→2A and Epa2A as well as the high-resolution structure of the Epa1A-T-antigen complex with its orthogonal binding mode for the GalNAc moiety challenge this notion and imply that the periphery formed by the loops L1 and L2 adds to the specificity of Epa adhesins.

In summary, our study suggests that variable presentation of Epa family members with diversified outer sub-sites at the cell surface of C. glabrata is one of the key elements for efficient and tissue-specific host invasion and could explain the clinical behavior of this human pathogenic fungus (23). Our structure-based insights that Candida–host interaction crucially depends on the conserved nature of an inner galactose-binding site of Epa adhesins may further contribute to the development of tailored antimycotics to combat this emerging pathogen.

Methods

Protein Production and Purification. Recombinant Epa1A and variants were generated in a thioredoxin and glutathione reductase-deficient Escherichia coli strain. Epa1A was purified by NiNTA-affinity and size-exclusion chromatography.

Phylogenetic Analysis. Phylogenetic analysis was performed by the neighbor-joining method, using Clustal X2.0 (24), COBALT (25), or a local copy of t-coffee (26) implemented with 3DCCoffee (27) for structure-based alignments. Preliminary targets were selected with the help of BLAST.

Crystalization and Structure Determination. Epa1A lactose cocrystals belonging to space group P2221 were obtained from ammonium sulfate-containing conditions. Epa1A-T-antigen cocrystals were obtained by soaking Ca²⁺-stripped Epa1A lactose cocrystals. The structure of Epa1A was then solved by molecular replacement and refined using REFMACS or PHENIX.

High-Throughput Glycan-Binding Assays. The Cfg glycan array consists of different groups of oligosaccharides that are presented by mammalian cells. Recombinant, fluorescently labeled Epa1A domains were applied to the Cfg array V4.1 chips at concentrations ranging from 1 μg/mL to 200 μg/mL. Chip surfaces were repeatedly washed; remaining fluorescence was measured and quantified.

Fluorescence Titrations. Fluorescence titrations of Epa1A and variants were performed against lactose and the T-antigen. Binding was followed at an emission wavelength of 346 nm by excitation of intrinsic tryptophan fluorescence at 295 nm. Fluorescence quench was recorded during titration and fitted using a one-site plus unspecific-binding model.

In Vivo Adhesion Assays. In vivo adhesion of Epa proteins to epithelial cells was determined using a nonadhesive S. cerevisiae strain presenting the different EpaA domains at the cell surface (Tables S3–S5). Adhesion assays were performed as previously described (28).

ACKNOWLEDGMENTS. The authors thank Tobias Klar and Alexander Popov for support at the European Synchrotron Radiation Facility, Grenoble (beamlines ID14-4 and ID23); Petra Gnaau and Lisa Ludewig for technical support; Uwe Linne for mass-spectrometric analyses; and David Smith and the Consortium for Functional Glycomics for performing in vitro screening of glycan specificity (Consortium for Functional Glycomics Request 2080). This research is supported by Grants ES15257, MO 825/1-4, and GRK 1216 from the Deutsche Forschungsgemeinschaft and by the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Excellenz (LOEWE) Center for Synthetic Microbiology (H.-U.M. and L.-O.E.).


Maestre-Reyna et al.